

# Radu Horaud

## List of Publications by Year in descending order

Source: <https://exaly.com/author-pdf/111070/publications.pdf>

Version: 2024-02-01

164  
papers

6,335  
citations

94381

37  
h-index

110317

64  
g-index

173  
all docs

173  
docs citations

173  
times ranked

4308  
citing authors

#	ARTICLE	IF	CITATIONS
1	Variational Inference and Learning of Piecewise Linear Dynamical Systems. IEEE Transactions on Neural Networks and Learning Systems, 2022, 33, 3753-3764.	7.2	3
2	The Impact of Removing Head Movements on Audio-Visual Speech Enhancement. , 2022, , .		1
3	Variational Bayesian Inference for Audio-Visual Tracking of Multiple Speakers. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43, 1761-1776.	9.7	27
4	Learning Visual Voice Activity Detection with an Automatically Annotated Dataset. , 2021, , .		5
5	Fullsubnet: A Full-Band and Sub-Band Fusion Model for Real-Time Single-Channel Speech Enhancement. , 2021, , .		88
6	Robust Face Frontalization For Visual Speech Recognition. , 2021, , .		2
7	A Comprehensive Analysis of Deep Regression. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42, 2065-2081.	9.7	148
8	How to Train Your Deep Multi-Object Tracker. , 2020, , .		120
9	A Recurrent Variational Autoencoder for Speech Enhancement. , 2020, , .		39
10	Audio-Visual Speech Enhancement Using Conditional Variational Auto-Encoders. IEEE/ACM Transactions on Audio Speech and Language Processing, 2020, 28, 1788-1800.	4.0	39
11	Neural network based reinforcement learning for audioâ€“visual gaze control in humanâ€“robot interaction. Pattern Recognition Letters, 2019, 118, 61-71.	2.6	33
12	Extended Gaze Following: Detecting Objects in Videos Beyond the Camera Field of View. , 2019, , .		14
13	Multichannel Online Dereverberation Based on Spectral Magnitude Inverse Filtering. IEEE/ACM Transactions on Audio Speech and Language Processing, 2019, 27, 1365-1377.	4.0	9
14	Online Localization and Tracking of Multiple Moving Speakers in Reverberant Environments. IEEE Journal on Selected Topics in Signal Processing, 2019, 13, 88-103.	7.3	29
15	Speech Enhancement with Variational Autoencoders and Alpha-stable Distributions. , 2019, , .		23
16	Tracking Multiple Audio Sources With the von Mises Distribution and Variational EM. IEEE Signal Processing Letters, 2019, 26, 798-802.	2.1	16
17	Semi-supervised Multichannel Speech Enhancement with Variational Autoencoders and Non-negative Matrix Factorization. , 2019, , .		41
18	Audio-Noise Power Spectral Density Estimation Using Long Short-Term Memory. IEEE Signal Processing Letters, 2019, 26, 918-922.	2.1	8

#	ARTICLE	IF	CITATIONS
19	Multichannel Speech Enhancement Based On Time-Frequency Masking Using Subband Long Short-Term Memory. , 2019, , .		27
20	Expectationâ€m maximisation for speech source separation using convolutive transfer function. CAAL Transactions on Intelligence Technology, 2019, 4, 47-53.	3.4	4
21	Multichannel Speech Separation and Enhancement Using the Convolutive Transfer Function. IEEE/ACM Transactions on Audio Speech and Language Processing, 2019, 27, 645-659.	4.0	15
22	Audio-Visual Variational Fusion for Multi-Person Tracking with Robots. , 2019, , .		4
23	Tracking Gaze and Visual Focus of Attention of People Involved in Social Interaction. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40, 2711-2724.	9.7	47
24	Plane-extraction from depth-data using a Gaussian mixture regression model. Pattern Recognition Letters, 2018, 110, 44-50.	2.6	12
25	Audio-Visual Speaker Diarization Based on Spatiotemporal Bayesian Fusion. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40, 1086-1099.	9.7	64
26	Joint Alignment of Multiple Point Sets with Batch and Incremental Expectation-Maximization. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40, 1397-1410.	9.7	93
27	Multisource Mint Using Convolutive Transfer Function. , 2018, , .		2
28	Deep Reinforcement Learning for Audio-Visual Gaze Control. , 2018, , .		6
29	Accounting for Room Acoustics in Audio-Visual Multi-Speaker Tracking. , 2018, , .		9
30	A VARIANCE MODELING FRAMEWORK BASED ON VARIATIONAL AUTOENCODERS FOR SPEECH ENHANCEMENT. , 2018, , .		49
31	Online Localization of Multiple Moving Speakers in Reverberant Environments. , 2018, , .		3
32	Multichannel Identification and Nonnegative Equalization for Dereverberation and Noise Reduction Based on Convolutive Transfer Function. IEEE/ACM Transactions on Audio Speech and Language Processing, 2018, 26, 1755-1768.	4.0	6
33	DeepGUM: Learning Deep Robust Regression with a Gaussian-Uniform Mixture Model. Lecture Notes in Computer Science, 2018, , 205-221.	1.0	10
34	Robust Head-Pose Estimation Based on Partially-Latent Mixture of Linear Regressions. IEEE Transactions on Image Processing, 2017, 26, 1428-1440.	6.0	62
35	Audio-visual tracking by density approximation in a sequential Bayesian filtering framework. , 2017, , .		6
36	Microphone array signal processing for robot audition. , 2017, , .		10

#	ARTICLE	IF	CITATIONS
37	Recognition of Group Activities in Videos Based on Single-and Two-Person Descriptors. , 2017, , .		2
38	Multiple-Speaker Localization Based on Direct-Path Features and Likelihood Maximization With Spatial Sparsity Regularization. IEEE/ACM Transactions on Audio Speech and Language Processing, 2017, 25, 1997-2012.	4.0	33
39	Switching Linear Inverse-Regression Model for Tracking Head Pose. , 2017, , .		7
40	Guest Editorial: Language in Vision. Computer Vision and Image Understanding, 2017, 163, 1-2.	3.0	0
41	An EM algorithm for joint source separation and diarisation of multichannel convolutive speech mixtures. , 2017, , .		6
42	Tracking a varying number of people with a visually-controlled robotic head. , 2017, , .		11
43	Audio source separation based on convolutive transfer function and frequency-domain lasso optimization. , 2017, , .		13
44	Deep Mixture of Linear Inverse Regressions Applied to Head-Pose Estimation. , 2017, , .		36
45	Exploiting the intermittency of speech for joint separation and diarization. , 2017, , .		5
46	Exploiting the Complementarity of Audio and Visual Data in Multi-speaker Tracking. , 2017, , .		15
47	An em algorithm for audio source separation based on the convolutive transfer function. , 2017, , .		10
48	Non-stationary noise power spectral density estimation based on regional statistics. , 2016, , .		12
49	Reverberant sound localization with a robot head based on direct-path relative transfer function. , 2016, , .		14
50	Simultaneous estimation of gaze direction and visual focus of attention for multi-person-to-robot interaction. , 2016, , .		7
51	EM Algorithms for Weighted-Data Clustering with Application to Audio-Visual Scene Analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38, 2402-2415.	9.7	86
52	An on-line variational Bayesian model for multi-person tracking from cluttered scenes. Computer Vision and Image Understanding, 2016, 153, 64-76.	3.0	33
53	An overview of depth cameras and range scanners based on time-of-flight technologies. Machine Vision and Applications, 2016, 27, 1005-1020.	1.7	190
54	An inverse-gamma source variance prior with factorized parameterization for audio source separation. , 2016, , .		6

#	ARTICLE	IF	CITATIONS
55	Estimation of the Direct-Path Relative Transfer Function for Supervised Sound-Source Localization. IEEE/ACM Transactions on Audio Speech and Language Processing, 2016, 24, 2171-2186.	4.0	42
56	Voice activity detection based on statistical likelihood ratio with adaptive thresholding. , 2016, , .		7
57	A Variational EM Algorithm for the Separation of Time-Varying Convolutional Audio Mixtures. IEEE/ACM Transactions on Audio Speech and Language Processing, 2016, 24, 1408-1423.	4.0	39
58	Tracking Multiple Persons Based on a Variational Bayesian Model. Lecture Notes in Computer Science, 2016, , 52-67.	1.0	35
59	A Distributed Architecture for Interacting with NAO. , 2015, , .		4
60	Cross-calibration of time-of-flight and colour cameras. Computer Vision and Image Understanding, 2015, 134, 105-115.	3.0	15
61	Local relative transfer function for sound source localization. , 2015, , .		7
62	A variational EM algorithm for the separation of moving sound sources. , 2015, , .		13
63	Tracking the Active Speaker Based on a Joint Audio-Visual Observation Model. , 2015, , .		19
64	Estimation of relative transfer function in the presence of stationary noise based on segmental power spectral density matrix subtraction. , 2015, , .		20
65	Head pose estimation via probabilistic high-dimensional regression. , 2015, , .		52
66	Robust Temporally Coherent Laplacian Protrusion Segmentation of 3D Articulated Bodies. International Journal of Computer Vision, 2015, 112, 43-70.	10.9	0
67	Fusion of Range and Stereo Data for High-Resolution Scene-Modeling. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37, 2178-2192.	9.7	29
68	Acoustic Space Learning for Sound-Source Separation and Localization on Binaural Manifolds. International Journal of Neural Systems, 2015, 25, 1440003.	3.2	58
69	Continuous Action Recognition Based on Sequence Alignment. International Journal of Computer Vision, 2015, 112, 90-114.	10.9	43
70	Vision-guided robot hearing. International Journal of Robotics Research, 2015, 34, 437-456.	5.8	28
71	Co-Localization of Audio Sources in Images Using Binaural Features and Locally-Linear Regression. IEEE/ACM Transactions on Audio Speech and Language Processing, 2015, 23, 718-731.	4.0	42
72	Hyper-Spectral Image Analysis With Partially Latent Regression and Spatial Markov Dependencies. IEEE Journal on Selected Topics in Signal Processing, 2015, 9, 1037-1048.	7.3	7

#	ARTICLE	IF	CITATIONS
73	High-dimensional regression with gaussian mixtures and partially-latent response variables. <i>Statistics and Computing</i> , 2015, 25, 893-911.	0.8	28
74	Continuous Gesture Recognition from Articulated Poses. <i>Lecture Notes in Computer Science</i> , 2015, , 595-607.	1.0	22
75	Audio-Visual Speech-Turn Detection and Tracking. <i>Lecture Notes in Computer Science</i> , 2015, , 143-151.	1.0	11
76	Skeletal Quads: Human Action Recognition Using Joint Quadruples. , 2014, , .		187
77	Audio-visual speaker localization via weighted clustering. , 2014, , .		10
78	Sound representation and classification benchmark for domestic robots. , 2014, , .		16
79	A Geometric Approach to Sound Source Localization from Time-Delay Estimates. <i>IEEE/ACM Transactions on Audio Speech and Language Processing</i> , 2014, 22, 1082-1095.	4.0	64
80	Automatic detection of calibration grids in time-of-flight images. <i>Computer Vision and Image Understanding</i> , 2014, 121, 108-118.	3.0	18
81	A Generative Model for the Joint Registration of Multiple Point Sets. <i>Lecture Notes in Computer Science</i> , 2014, , 109-122.	1.0	44
82	RAVEL: an annotated corpus for training robots with audiovisual abilities. <i>Journal on Multimodal User Interfaces</i> , 2013, 7, 79-91.	2.0	27
83	Variational EM for binaural sound-source separation and localization. , 2013, , .		19
84	The geometry of sound-source localization using non-coplanar microphone arrays. , 2013, , .		0
85	From Video Matching to Video Grounding. , 2013, , .		2
86	Time-of-Flight Cameras. <i>SpringerBriefs in Computer Science</i> , 2013, , .	0.2	111
87	Characterization of Time-of-Flight Data. <i>SpringerBriefs in Computer Science</i> , 2013, , 1-28.	0.2	2
88	Real-time visuomotor update of an active binocular head. <i>Autonomous Robots</i> , 2013, 34, 35-45.	3.2	6
89	Alignment of binocular-binaural data using a moving audio-visual target. , 2013, , .		4
90	Benchmarking methods for audio-visual recognition using tiny training sets. , 2013, , .		0

#	ARTICLE	IF	CITATIONS
91	Active-speaker detection and localization with microphones and cameras embedded into a robotic head. , 2013, , .		20
92	A Mixed Time-of-Flight and Stereoscopic Camera System. SpringerBriefs in Computer Science, 2013, , 77-96.	0.2	0
93	Calibration of Time-of-Flight Cameras. SpringerBriefs in Computer Science, 2013, , 45-58.	0.2	4
94	The cocktail party robot. , 2012, , .		27
95	Audio-visual robot command recognition. , 2012, , .		1
96	High-resolution depth maps based on TOF-stereo fusion. , 2012, , .		42
97	Sound-event recognition with a companion humanoid. , 2012, , .		18
98	2D sound-source localization on the binaural manifold. , 2012, , .		22
99	Online multimodal speaker detection for humanoid robots. , 2012, , .		7
100	Action Recognition Robust to Background Clutter by Using Stereo Vision. Lecture Notes in Computer Science, 2012, , 332-341.	1.0	7
101	Keypoints and Local Descriptors of Scalar Functions on 2D Manifolds. International Journal of Computer Vision, 2012, 100, 78-98.	10.9	96
102	Scale-space representation of scalar functions on 2D manifolds. , 2011, , .		0
103	Topology-Adaptive Mesh Deformation for Surface Evolution, Morphing, and Multiview Reconstruction. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33, 823-837.	9.7	38
104	Rigid and Articulated Point Registration with Expectation Conditional Maximization. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33, 587-602.	9.7	170
105	Finding audio-visual events in informal social gatherings. , 2011, , .		16
106	A Differential Model of the Complex Cell. Neural Computation, 2011, 23, 2324-2357.	1.3	5
107	Conjugate Mixture Models for Clustering Multimodal Data. Neural Computation, 2011, 23, 517-557.	1.3	21
108	Scene flow estimation by growing correspondence seeds. , 2011, , .		59

#	ARTICLE	IF	CITATIONS
109	Projective alignment of range and parallax data. , 2011, , .		16
110	Topologically-robust 3D shape matching based on diffusion geometry and seed growing. , 2011, , .		41
111	An Unsupervised Framework for Action Recognition Using Actemes. Lecture Notes in Computer Science, 2011, , 592-605.	1.0	2
112	Disparity and normal estimation through alternating maximization. , 2010, , .		0
113	Cyclorotation Models for Eyes and Cameras. IEEE Transactions on Systems, Man, and Cybernetics, 2010, 40, 151-161.	5.5	9
114	Surface feature detection and description with applications to mesh matching. , 2009, , .		282
115	Robust Factorization Methods Using a Gaussian/Uniform Mixture Model. International Journal of Computer Vision, 2009, 81, 240-258.	10.9	17
116	Human Motion Tracking by Registering an Articulated Surface to 3D Points and Normals. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31, 158-163.	9.7	39
117	Inexact Matching of Large and Sparse Graphs Using Laplacian Eigenvectors. Lecture Notes in Computer Science, 2009, , 144-153.	1.0	27
118	Human Motion Tracking with a Kinematic Parameterization of Extremal Contours. International Journal of Computer Vision, 2008, 79, 247-269.	10.9	42
119	Articulated shape matching using Laplacian eigenfunctions and unsupervised point registration. , 2008, , .		116
120	Cooperative disparity and object boundary estimation. , 2008, , .		1
121	Cyclopean geometry of binocular vision. Journal of the Optical Society of America A: Optics and Image Science, and Vision, 2008, 25, 2357.	0.8	28
122	Coherent Laplacian 3-D protrusion segmentation. , 2008, , .		4
123	Temporal Surface Tracking Using Mesh Evolution. Lecture Notes in Computer Science, 2008, , 30-43.	1.0	34
124	The CAVA corpus. , 2008, , .		12
125	Detection and localization of 3d audio-visual objects using unsupervised clustering. , 2008, , .		6
126	Robust Spectral 3D-Bodypart Segmentation Along Time. , 2007, , 196-211.		1

#	ARTICLE	IF	CITATIONS
127	Articulated Shape Matching Using Locally Linear Embedding and Orthogonal Alignment. , 2007, , .		11
128	Articulated-Body Tracking Through Anisotropic Edge Detection. Lecture Notes in Computer Science, 2007, , 86-99.	1.0	3
129	Articulated Shape Matching by Robust Alignment of Embedded Representations. , 2007, , .		10
130	Patterns of Binocular Disparity for a Fixating Observer. , 2007, , 308-317.		5
131	TransforMesh : A Topology-Adaptive Mesh-Based Approach to Surface Evolution. , 2007, , 166-175.		66
132	Camera cooperation for achieving visual attention. Machine Vision and Applications, 2006, 16, 1-2.	1.7	42
133	Multiple Camera Calibration Using Robust Perspective Factorization. , 2006, , .		11
134	Tracking with the Kinematics of Extremal Contours. Lecture Notes in Computer Science, 2006, , 664-673.	1.0	4
135	Image interpolation for virtual sports scenarios. Machine Vision and Applications, 2005, 16, 236-245.	1.7	2
136	Image matching with scale adjustment. Computer Vision and Image Understanding, 2004, 93, 175-194.	3.0	45
137	Motion Panoramas. Computer Animation and Virtual Worlds, 2004, 15, 501-517.	0.7	33
138	Hand Motion from 3D Point Trajectories and a Smooth Surface Model. Lecture Notes in Computer Science, 2004, , 495-507.	1.0	45
139	Special Issue on Visual Analysis of Human Movement. International Journal of Robotics Research, 2003, 22, 357-357.	5.8	0
140	Visual Servoing/Tracking Using Central Catadioptric Images. , 2003, , 245-254.		34
141	Visual Servoing from Lines. International Journal of Robotics Research, 2002, 21, 679-699.	5.8	83
142	Multiple-Camera Tracking of Rigid Objects. International Journal of Robotics Research, 2002, 21, 97-113.	5.8	22
143	Selection of Marr Prize and Honourable Mentions at ICCV'01. International Journal of Computer Vision, 2002, 48, 7-8.	10.9	0
144	Motionâ€Eg motion Discrimination and Motion Segmentation from Image-Pair Streams. Computer Vision and Image Understanding, 2000, 78, 53-68.	3.0	18

#	ARTICLE	IF	CITATIONS
145	Visual Servoing Using Projective Kinematics. , 2000, , 97-104.		1
146	Iterative Pose Computation from Line Correspondences. Computer Vision and Image Understanding, 1999, 73, 137-144.	3.0	74
147	Finding the Collineation between Two Projective Reconstructions. Computer Vision and Image Understanding, 1999, 75, 260-268.	3.0	20
148	Visually guided object grasping. IEEE Transactions on Automation Science and Engineering, 1998, 14, 525-532.	2.4	107
149	Simultaneous robot-world and hand-eye calibration. IEEE Transactions on Automation Science and Engineering, 1998, 14, 617-622.	2.4	223
150	<title>Terminal air-to-ground missile guidance by infrared seeker</title>. , 1997, , .		0
151	Integrating grasp planning and visual servoing for automatic grasping. , 1997, , 71-82.		6
152	Stability of lagrangian duality for nonconvex quadratic programming. Solution methods and applications in computer vision. ESAIM: Mathematical Modelling and Numerical Analysis, 1997, 31, 57-90.	0.8	3
153	Object Pose: The Link between Weak Perspective, Paraperspective, and Full Perspective. International Journal of Computer Vision, 1997, 22, 173-189.	10.9	115
154	Introduction to the special volume on computer vision. Artificial Intelligence, 1995, 78, 1-3.	3.9	2
155	Polyhedral object recognition by indexing. Pattern Recognition, 1995, 28, 1855-1870.	5.1	42
156	Object pose from 2-D to 3-D point and line correspondences. International Journal of Computer Vision, 1995, 15, 225-243.	10.9	129
157	Smooth curve extraction by mean field annealing. Annals of Mathematics and Artificial Intelligence, 1995, 13, 281-300.	0.9	0
158	Hand-Eye Calibration. International Journal of Robotics Research, 1995, 14, 195-210.	5.8	430
159	Figure-ground discrimination: a combinatorial optimization approach. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1993, 15, 899-914.	9.7	124
160	On single-scanline camera calibration. IEEE Transactions on Automation Science and Engineering, 1993, 9, 71-75.	2.4	53
161	Finding geometric and relational structures in an image. Lecture Notes in Computer Science, 1990, , 374-384.	1.0	63
162	An analytic solution for the perspective 4-point problem. Computer Vision, Graphics, and Image Processing, 1989, 47, 33-44.	1.1	292

#	ARTICLE	IF	CITATIONS
163	On the geometric interpretation of image contours. Artificial Intelligence, 1988, 37, 333-353.	3.9	61
164	New Methods for Matching 3-D Objects with Single Perspective Views. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1987, PAMI-9, 401-412.	9.7	132